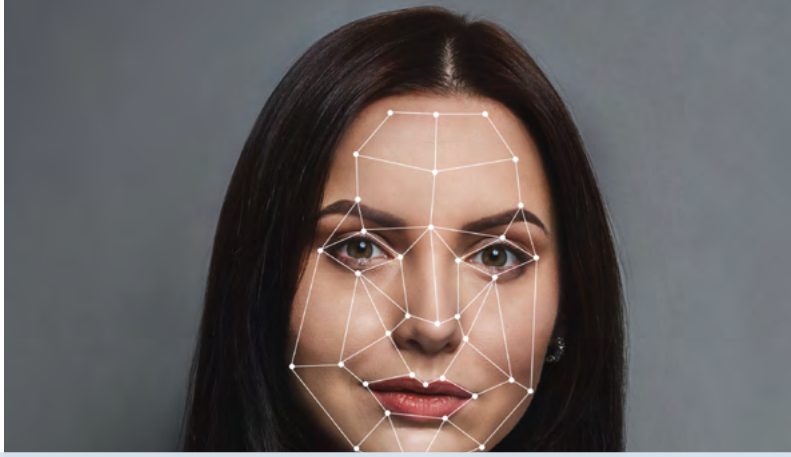
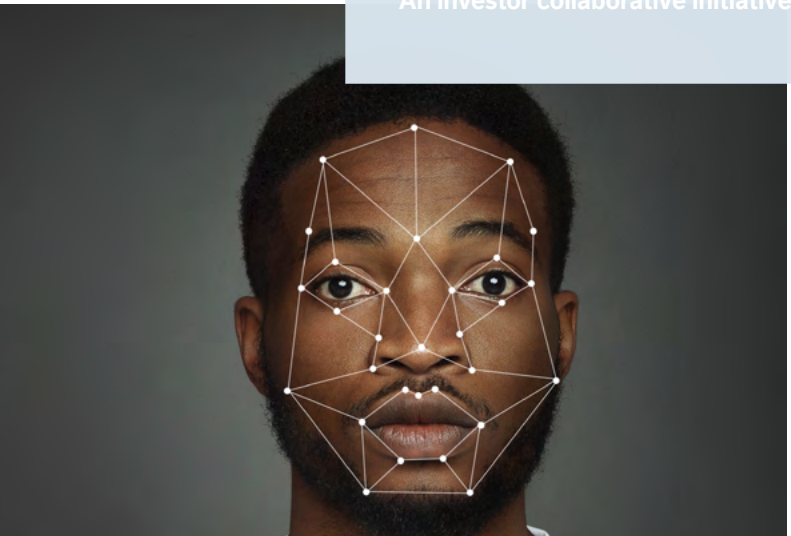



September 2022

Investor Engagement on Human Rights Risks of Facial Recognition Technology: Interim Report

An investor collaborative initiative launched and coordinated by Candriam




*These dialogues were conducted by, and best practice recommendations offered by, an ad hoc working group of investor signatories of the **Investor Statement on Facial Recognition**.*

This document was approved in September, 2022 by the working group members whose names appear on pages 22-23.

Table of contents

Investor Engagement on Human Rights Risks of Facial Recognition Technology: Interim Report

Governments, law enforcement agencies and corporates are rushing to adopt Facial Recognition Technology (FRT) with a hope of increasing security, as well as improving efficiency and customer experience. Investors are highly interested. FRT, along with other Artificial Intelligence (AI) functions, is a promising technology. Concurrently, its rapid deployment has raised major concerns about the impact on privacy, data protection and civil liberties.

It is estimated that 1 billion surveillance cameras were in operation around the globe by the end of 2021 — one camera for every 8 citizens of the world. Yet, the way this technology is being used and designed carries considerable risks to basic human rights as well as far reaching social implications.

Facial Recognition in Context: Investors are Interested

FRT generated opportunities, including investment opportunities, as well as grave risks. As responsible investors we set out to understand, alert, discuss and mitigate the risks linked to FRT. We spoke to experts, academics, and journalists, summarizing the findings in a white paper.

Investors are interested. By June 2022, a group of 55 investors representing over $5 trillion in assets under management had signed a statement to alert companies to the serious risks posed by this technology.

What causes these concerns?

- The **racial and gender biases** observed in these systems
- The **questionable accuracy** and lack of public testing of most systems in use
- Possible **privacy** or legal violations in the sourcing of photos for databases
- **Misuse** by some governments, law enforcement agencies or others

The statement signed by this group of investors sets out a list of expectations for companies developing and/or using Facial Recognition Technology.

As signatories, we expect companies to:

- Demonstrate that their technology is constantly monitored to detect algorithmic biases, particularly with respect to race, gender, or age.

- Disclose the accuracy of their technology, measured by a recognized and relevant scientific assessment institution.
- Disclose the sources of their image databases.
- Demonstrate proper due diligence of clients before making the technology available to them.
- Demonstrate that effective grievance mechanisms are in place to enable victims to report consequences and to access remedies.

This statement was sent to over 60 prominent companies involved in the technology.

A group of 20 investor signatories engaged with 15 companies to gain a better understanding of their policies and procedures surrounding Facial Recognition Technology.

Figure 1: Significant Developments in FRT Since the Launch of the FRT investors statement

MAY 2021

The European Commission issues a draft regulation on AI. The proposal defines four categories of risk to the public: Unacceptable Risk AI systems, High Risk AI, Limited Risk AI, and Minimal Risk AI systems. Most FRT systems will be qualified as 'High Risk AI'.

OCTOBER 2021

In anticipation of the negotiations around the EU AI Act, the European Parliament adopted a non-binding resolution calling for a ban on police use of facial recognition technology in public places, and on predictive policing (profiling controversial practice before any crime is committed), and requested a ban on private facial recognition databases.

NOVEMBER 2021

Facebook discontinues its facial recognition feature and deletes 1 billion 'face prints'. The announcement mentions three important points aligned with our investor statement goals: an admission that a safe use of FRT can not yet be guaranteed, that FRT has strong repercussions on society, and that regulation is rapidly needed.

MARCH 2022

Ukraine implements Facial Recogntion software to identify Russian combat assailants and operatives, as well as to identify combatant casualties. The technology is also being used to counter misinformation.

MARCH 2022

China issues an ambitious regulation on Artificial Intelligence becoming one of the first nations to regulate AI algorithms. The regulation defines 'Trustworthy' Facial Recognition applications and protection plans.

JUNE 2022

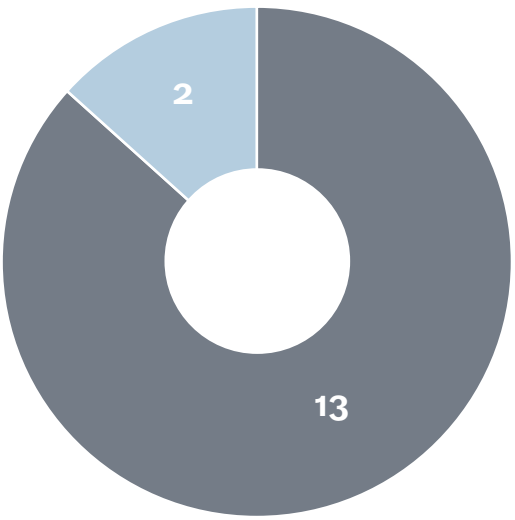
Microsoft retires Facial Analysis Capabilities that purport to infer emotional states and identity sensitive attributes such as gender, age, smile, facial hair, hair, and makeup. The company will also restrict access to its facial recognition services.

Scope and Methodology for Engagement

Scope

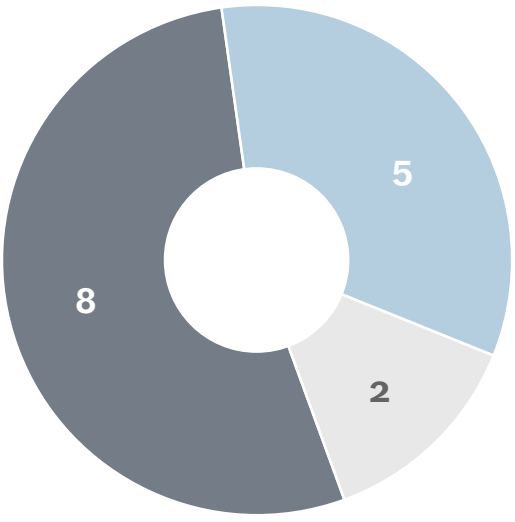
To choose companies for engagement, we considered the level of involvement in FRT, their importance in the FRT market and value chain, the relative importance of FRT to each company's business, and existing public statements by each company. Share ownership by investor(s) in this group was not a part of the determination. We selected both public and private firms.

Figure 2: Profile of the Companies Contacted



Ownership

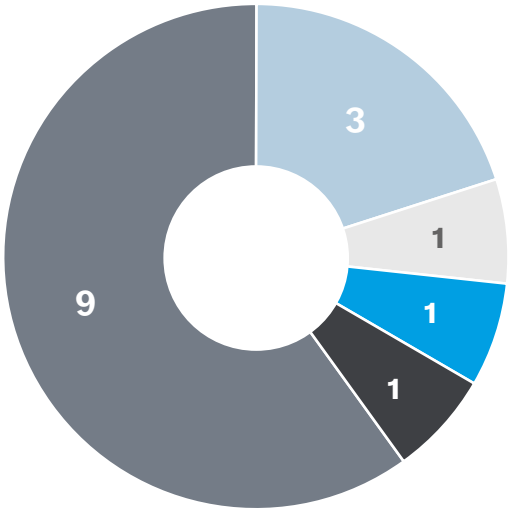
Public	13
Private	2

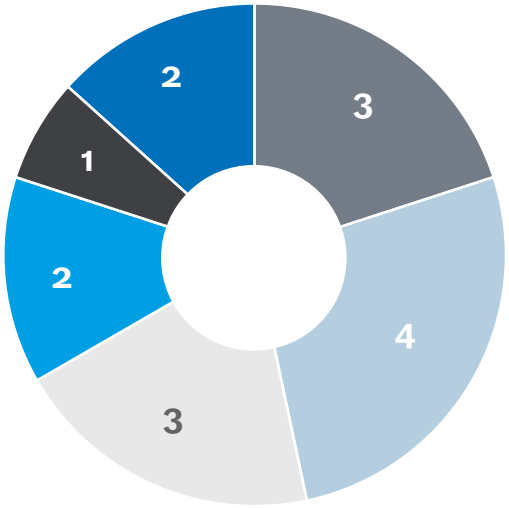


Region

North America	8
Asia	5
Europe	2

The sample includes a wide variety of market sizes and sectors. However, it is often difficult to determine when / which product or services are linked to FRT. At the bottom of this page, we show a list of the product and / or services in which the 15 companies we engaged are or were involved.





Sector

Software	9
Semiconductors	3
Internet Platform	1
Internet Retailing	1
Telecommunications	1

Market Capitalization of listed companies

Below 30bln$	3
Between $30bln and $100bln	4
Between $100bln and $200bln	3
Between $200bln and $1,000bln	2
Above $1,000bln	1
Unlisted	2

Products and services offered by the companies engaged:

• FRT-ready microprocessors / semiconductors for smartphones, cameras, computers	• Facial Recognition Categorisation Software
• Biometric database provider	• Facial Recognition Surveillance Software
• Facial Recognition Identification Software	• Biometric Search Engine
• Facial Recognition Authorisation Software	• Biometric object detection Software
• Facial Recognition Hardware Portal / Access Control	

Methodology for Engagement

We began our research by speaking to industry experts, academics, journalists. The research findings are detailed in Candriam's white paper, Investor Guidance on Facial Recognition. This, along with further exchanges with experts during the engagement period, enabled us to define a framework to analyse the FRT practices of companies.

It was important to develop this framework before speaking with companies so that we could compare their responses, and hopefully aggregate some of their best approaches. This framework consists of questions under six major themes:

1. Definition of Facial Recognition Product/Service offered by the engaged company
2. Company Policy
3. Governance Structure and procedures of the company
4. Product Design focus
5. Product Use focus
6. Remedy channels focus

The Investor Statement was sent to sixty companies in June, 2021. Fifteen of these companies were sent invitations for dialogues during the fourth quarter of 2021, and the conversations took place between December 2021 and May 2022. Typical engagement meetings included three or more investor representatives and relevant officers such as those in charge of ethics, sustainability, legal, compliance, and technology from the company side. Some exchanges required follow-up, either through a second meeting, or by email.

Engagement Results and Emerging Best Practices

Looking for a Quick-Start Guide? Dialogues can share information in several directions, among multiple stakeholders – and from them, best practices become better practices. A fascinating example is that accuracy is improved by combining the algorithm with a human check. But if you don't have time to read the results of the dialogues – then may we suggest the 'Best Practice' sections in the blue boxes?

Constructive Conversations

A number of companies were eager to speak to us. They were keen to share their best practices and felt that the industry was overdue some form of guard rail, or regulation, that would impose better and more consistent procedures.

Others were less open, responded via email, or did not respond at all. Unsurprisingly, these companies showed the weakest practices.

We also spoke with companies which described themselves as being at the start of their learning process. These companies were seeking to improve their procedures, and were extremely interested in our initiative. They hope the information sharing will enable them to identify best practices and adapt them to their own businesses.

If we were to make any generalizations, we might conclude that the closer an organization is to the algorithm, the greater awareness the company has. Thus, in this sample, the software industry seemed to display the strongest practices and respect for human rights. Conversely, despite developing chips specifically for Facial Recognition systems, semiconductor companies showed little interest for the potential mis-use of their devices further down the chain. Is it possible a company considers itself to have less responsibility when it is further away from the end use of the product?

Policy and Ethics: Dialogue Results

Artificial Intelligence, Facial Recognition Technology, and Human Rights

Every company we spoke to, and every company we analysed – an even greater number — for this initiative has some form of Human Rights Policy. It is a pre-requisite today to express some sort of respect for the protection of human rights. Most companies subscribe to the UN Guiding Principles on Business and Human Rights. Yet very few explicitly mentioned the risks of Artificial Intelligence (AI), and none mentioned Facial Recognition Technology in their policies.

Public Acknowledgement

Many companies have written papers on AI and FRT. Some companies, such as Microsoft and Thales, have made their work public. These demonstrate an awareness of the risks associated with AI technologies in general, and Facial Recognition in particular.

Legal, Yes — but Ethical?

Some companies, particularly those in the semi-conductor industry, indicate that selling to authorised customers in authorised countries satisfies their obligation. These companies were not concerned by the potential abuse or misuse of their products or services further down the line. They had little in place with respect to ethics policy, human rights impact assessment or related due diligence.

Facial Recognition and Law Enforcement

Based on the exchanges we had, the year 2020 was an important milestone. The Black Lives Matter movement, the use of FRT software on crowds of protesters, and the potential abuse by law enforcement triggered deeper thinking at some large US technology firms. These paused their sales of FRT to law enforcement agencies. The ensuring debate prompted a strong civil liberty protection movement to prevent or restrict the use of the technology by police forces.

Categorisation

FRT is capable of several tasks. It can recognise whether a picture contains a face, identify or authenticate the face, and some algorithms attempt to categorise the face, including attempts to predict sexual orientation or people likely to commit terrorist acts. Such function can define the gender, age, race or even emotions of the face it is analysing. With or without errors, these functions can lead to huge human rights violations. Some companies we exchanged with in the industry have decided that these functions are not ethical and should not be deployed. Microsoft recently updated its Responsible AI Standard to reflect its concerns over classification of sensitive facial attributes, as well as the responsible use of its facial analysis tools overall.

Policy and Ethics: Best Practices

- Companies involved in AI and FRT should mention these technologies in their Human Rights Policies. They should acknowledge and assess the risks these technologies create and establish vigilant monitoring of these risks, and conduct human rights impact assessments.
- Companies should communicate, through white papers, their position concerning these technologies and describe their approach to mitigating human rights risks.
- As technology is advancing much more rapidly than guardrails or regulation, companies should demonstrate publicly that they constantly seek to adopt the highest ethical standard and specify their lobbying practices.
- Companies should take broader responsibility when there is clear evidence of risks. For example, companies should pause sales to law enforcement agencies of products, services, and databases until adoption of strong regulation and oversight.
- Companies should restrict designing, using, or offering categorisation functionalities.
- Companies should conduct due diligence on customers to evaluate the potential risks of intended use cases and put in place clear guidelines and terms of use.

Case Study: Microsoft

Through its Azure Face service Microsoft offers its customers APIs that can detect, recognize, and analyze human faces in images.

Microsoft offers a facial recognition product which customers can embed into their own software applications to verify identity or control access. The company provided a very detailed explanation of its approach to both Facial Recognition, and Responsible AI in general.

Our conversation(s) took place with a key officer in their ethics process. The company has spent years defining and refining high standards of responsibility. In 2018 Microsoft released a set of six guiding principles for its Facial Recognition systems, which we believe can inspire other participants in the technology.

1. ***Fairness.*** *We [Microsoft] will work to develop and deploy facial recognition technology in a manner that strives to treat all people fairly.*
2. ***Transparency*** *– We will document and clearly communicate the capabilities and limitations of facial recognition technology.*
3. ***Accountability****. We will encourage and help our customers to deploy facial recognition technology in a manner that ensures an appropriate level of human control for uses that may affect people in consequential ways.*
4. ***Non-discrimination****. We will prohibit in our terms of service the use of facial recognition technology to engage in unlawful discrimination.*
5. ***Notice and consent****. We will encourage private sector customers to provide notice and secure consent for the deployment of facial recognition technology.*
6. ***Lawful surveillance****. We will advocate for safeguards for people's democratic freedoms in law enforcement surveillance scenarios and will not deploy facial recognition technology in scenarios that we believe will put these freedoms at risk.*

Microsoft regularly reviews and updates its internal guidelines for designing, building, and testing AI responsibly. The company publicly released its most recent updated Responsible AI Standard (June, 2022).

- In its Responsible AI Standard Microsoft details how they implement the above 6 principles. One of the important goals defined in this document is to ensure **Human Oversight**.

- Microsoft carries out extensive impact assessments to identify potential adverse impact of sensitive AI technologies. The company uses a '**Responsible AI Impact Assessment**' to define how AI systems can affect 'people, organisations and society'. This template is a public document and is shared with the rest of the industry.

- The company also carries out **oversight** to identify potential misuse and abuse by its customers of AI and FRT systems. - Through various channels, Microsoft **engages** with customers, public officials, technologists, academics, civil society groups, and multi-stakeholder organizations such as the 'Partnership on AI' to keep up with industry practices and challenges. Microsoft is also an **advocate for strong regulation** in the field of AI and Facial Recognition Technology.

Product Definition and Design: Dialogue Results

Databases

Every company that develops Facial Recognition Technology (and other forms of biometric identification systems) utilizes a training database to teach its algorithms. The greater the database, the more accurate the system. But how were these databases collected? Did those whose biometric data were collected give *explicit* consent? The companies that answered this question all said they used, rented, or purchased databases that were free of consent.

Accuracy

Most, but not all, of the companies we spoke with have submitted their algorithms to independent public testing. The main body for this testing is the U.S. National Institute of Standards and Technology (NIST). A few launched 'hackathons' to challenge their algorithm for potential biases or inaccuracies.

Overlaying Human Judgement on top of Machine Decisions

A NIST survey from 2018 found that the optimal facial recognition testing results were obtained not by human, nor by machine, but by the overlay of human decision on top of machine systems. Interestingly, two prominent companies we spoke to have already put in place robust procedures with mandatory human oversight of their machine systems. In one specific example, a company manufacturing border control e-passport gates, explained that their equipment can only be operated if each five gates are being controlled by one officer. This allows for human supervision, as well as filtering, so that a particular percentage of the least certain identifications are double-checked by human judgement.

Encrypted biometric Data – No Undue Recording of Data

Companies operating authentication systems (such as building access control) told us their systems do not allow for the recording of biometric data once the authentication was carried out. After authentication, the biometric data is immediately deleted from the system.

Product Definition and Design: Best Practices

- Companies should publish the source, size and characteristics of any biometric database they have gathered, used, purchased or rented for the design of an FRT product or service.
- Companies engaging in FRT should submit their algorithms to independent public testing and make the results publicly available.
- Companies should systematically ensure that at least one human validates the results of FRT algorithms in the case of material/serious/non-trivial uses (access to a border is different from accessing an office).
- Companies operating FRT systems where data is being collected should constantly destroy non-essential biometric data/ records after use.

Case Study:
Motorola Solutions

Motorola Solutions offers various security alert features that utilize facial recognition technology as part of its video security solutions offerings. We believe some of Motorola Solutions' policies and procedures, presented below, offer insight for other providers and users of FRT, and for investors and other stakeholders.

- **Governance** — The company established an internal, cross-functional Motorola Solutions Advisory Committee (MTAC) on ethics, limitations, and implications of specific product technologies. This committee also engages with external stakeholders to obtain objective inputs to guide decisions on sensitive technologies. Before deploying AI technology, including FRT, Motorola Solutions weighs the benefits against potential harmful risks.

- **Human in the Loop** — Motorola Solutions states that AI should be used to assist and accelerate human decision-making, not replace it. Its AI systems are advisory in nature and are not intended to take consequential actions on their own. Using AI-generated guidance, these systems are designed to help appropriate users make better decisions faster.
- **Authentication** — Two-factor authentication is built into the technlogy. This step encourages more than one human to review the results before committing to an actionable identification. It can be triggered through an automatic audit process. An individual can trigger a peer review to request another, independent verification of the results. The goal is to enable greater trust, by allowing for human intervention in the technology.
- **Disciplined Innovation** — The company applies tested, characterized, and trained AI and machine learning technologies. This demonstrates that Motorola Systems is making a commitment to solutions that are far less likely to behave unexpectedly in the field.
- **Oversight** — In addition to the MTAC, Motorola Solutions consults with objective third parties to provide an outside-in point of view to guide its decisions.
- **Policy** — Motorola has a written policy specifically for data rights and ethics, and makes it publicly available.

Governance, Marketing, and Product Use: Dialogue Results

Governance of Human Rights / Algorithmic Risk

The strongest companies had at least one board member, *with appropriate experience*, who is responsible for oversight of human rights risks. Some companies with strong human rights governance have an ethics committee composed of experienced officers reporting to the board. Regular engagement with outside expert counsel, think tanks or NGOs is also good practice.

Client Due Diligence

Few companies consider the risk potential which can arise from harmful use of their hardware or software. It was evident with a semiconductor company, for example, that the company was unconcerned with the end use of its product. Their view was that strict obedience to rules and regulations – the letter of the law – is sufficient to guarantee 'the right thing'.

Systems Alerts and Audit Trails

Some companies build alerts into their software products to detect and report abusive or abnormal behaviour. For example, the identification of a disproportionate number of images of a certain category of the population by a police officer using Facial Recognition would generate an alert. Some companies actually propose to log any use of FRT algorithms at all times, so that efficient oversight and monitoring can be applied.

Client Training

Most system developers deploying FRT for outside customers offer some type of training. Although we had only a handful of discussions with corporates concerning training, we were alarmed at how little training was offered, especially on the topic of protecting the rights of those subject to facial recognition. Information sessions ranged from a few hours to two days. Training covered the use of the systems and software, but offered no warnings or real understanding on the potential use and mis-use of the technology.

Opt-Out

We asked most companies if their systems offered an opt-out feature. In Europe, the General Data Protection Regulation (GDPR) requires such functionality, yet even in Europe, many uses of FRT still apparently do not comply. And, when available, many public uses of these systems are accompanied by warning signs that are, to say the least, 'discrete'

Facilitated access to information is of high importance for all groups of users, but perhaps even more so for children and other vulnerable categories of users. The topic of child protection came up only twice in our conversations with companies.

Governance, Marketing, and Product Use: Best Practice

- Companies involved in AI and more specifically in FRT should have strong human rights governance in place, including:
  - Board members with recognized human rights expertise,
  - An ethics committee reporting to the board composed of internal and external members (ideally representative of a least two recognised NGOs/think tanks should sit on this committee), and
  - A human rights ombudsman in every business unit.
- Every use of FRT should be disclosed, while offering an opt-out option for those subject to the technology. An alternative source of identification or authentication should always be available – and clearly offered — to those who do not wish to have their biometric data analysed by an algorithm.
- Companies should perform extensive risk-level-based due diligence *of their clients* to prevent the mis-use or abuse of their technology. While the technology is deployed companies should ensure extensive and periodic client training is conducted, with special attention to the potential human rights violations.
- Companies should implement strong monitoring and oversight procedures, including alert systems that warn management of potential mis-use of the technology, as well as a process for discontinuing client relationships where human rights abuse is suspected. Companies should also establish grievance mechanisms.

Case Study: Thales

Thales offers identification systems, such as the passport gates at Paris Charles de Gaulle Airport. The company has strong procedures in place to ensure its systems are ethical, safe, and ensure confidentiality. Thales communicates about its work on the ethics of AI in general and of FRT in particular.

- Thales designs its FRT systems with strict ethical principles in mind under the Thales TrUE AI approach, an acronym for Transparent, Understandable and Ethical artificial intelligence.
  - ***Transparent AI***, where users can access the data used to arrive at a conclusion,
  - ***Understandable AI***, that can explain and justify the results,
  - ***Ethical AI***, that follows objective standards, protocols, laws, and human rights.

Picture credit: Thales.

- Thales' FRT solutions are designed to follow essential rules.

Confidentiality and consent.

- ***Transparency***.
- ***Precision and reliability***.
- ***Security***.
- ***Ethics & Compliance***.
- ***Accountability***.

As a result, Thales' systems offer some interesting features.

- Verification is carried out in a 'closed environment'. That is, the system checks only the biometric data encrypted in the travel document against the person in front of the camera, not against a database. The environment is optimized for facial recognition precision, in that the person being photographed is not moving, and the camera angle is also optimal.

- The data is destroyed after the passenger has departed.

- An **age restriction** is imposed on the border control passport gates. Children under 15 are not allowed to use the system.

- Once Thales defines a biometric facial recognition product as 'Sensitive', the company subjects it to extra safety governance.

- Thales submits its systems to the US National Institute of Standards and Technology, and discloses publicly the results of this independent organization.

- Thales reinforced its board committee in charge of CSR topics with two directors with Corporate Social Responsibility expertise.

Shaping the Future?

We have mapped company practices and presented those we believe represent today's best.

Even with these, some of the expectations we had listed in our initial FRT statement, such as the existence of effective grievance mechanisms, remain untouched.

The next step is to discuss with each company how identified best practices could be implemented in their own organization.
The outcome of this second campaign of dialogues should be released in 2023.

In parallel, our initiative has also attracted interest from two major global organisations. By sharing with the World Economic Forum's Responsible AI Framework and collaborating with the United Nations B-Tech Project, we can all help promote best practices with our partners.

For us, as Responsible Investors, collecting and sharing information, contributing to enhancement of practices is a way to reduce uncertainties associated with this promising technology. Engagement and communication of expectations can aid our understanding of company practices to enhance our analysis. Our role is also to support a better and effective inclusion of the interest of all stakeholders in its development and use. It is also an opportunity to pave the way for assessment of other AI technologies which are increasingly prevalent in our society, while we continue to struggle to clarify limits.

Investors in the Dialogue Working Group:

A collaborative initiative launched by Candriam

CANDRIAM
A NEW YORK LIFE INVESTMENTS COMPANY

æquo

Assenagon

Aviva Investors

BMO Global Asset Management*

Boston Common Asset Management

Columbia Threadneedle Investments

Degroof Petercam Asset Management

Domini

EdenTree Investment Management

Ethos

*BMO Global Asset Management (BMO GAM) was a participant in the initiative in 2021. The European division of BMO GAM (BMP GAM EMEA) was acquired by Columbia Threadneedle Investments in November 2021 and continued this participation. BMO GAM EMEA rebranded fully to Columbia Threadneedle Investments in July 2022.

Impact AM

Mercy Investments

NEI Investments

New Zealand Superannuation Fund

Öhman

Railpen

Robeco

Share

Sycomore Asset Management

Vancity Investment Management